as MIAME, are bound to face considerable challenges in the future. In fact, it is possible that an as yet unknown or 'fringe' technology could displace all current oligonucleotide and cDNA-based platforms.

John Weinstein (National Cancer Institute, NIH, Bethesda, MD, USA) pointed out, in a round-table discussion, that it might be 'too soon to standardize' microarray databases and warned that the community runs the risk of establishing a standard that actually 'restricts rather than enables.' After all, there has not yet even been a standardization of the most fundamental unit of a transcriptional profiling experiment, gene annotations.

Future directions

This conference provided a current perspective on the field of transcriptional profiling. It is clear that the field has moved well beyond the era of pure technology development, where success could be measured simply by the production of a functional microarray. We now find ourselves in the luxurious position of contemplating proper experimental design, choice of statistical methods and database construction – a sure sign of progress. Nonetheless, the challenges of data integration and compilation are readily apparent. Just as the utility of the first microbial genome sequence has increased by the sequencing of other genomes, the overall utility of any one series of microarray experiments will scale over time with increased data for comparison. Identifying enabling methods of standardization across technology platforms, processing sites and biological systems will be essential for the future success of this endeavor.

References

1 Box, G.E.P. *et al.* (1978) *Statistics for experimenters*. (1st edn), John Wiley and Sons

2 Golub, T.R. *et al.* (1999) Molecular classification of cancer: class discovery and class prediction by gene expression profiling. *Science* 286, 531–537

3 Brazma, A. *et al.* (2001) Minimum information about a microarray experiment (MIAME) – toward standards for microarray data. *Nat. Genet.* 29, 365–367

Chemical genomics: discovery of disease genes and drugs

Moulay H. Alaoui-Ismaili, Peter T. Lomedico and Satish Jindal*, NeoGenesis Pharmaceuticals, 840 Memorial Drive, Cambridge, MA 02139, USA, *tel: +1 617 868 1500, fax: +1 617 868 1515, e-mail: satjin@neogenesis.com

Chemical genomics, a popular and efficient technology for exploiting genomic information, was the subject of a recent Cambridge Healthtech Institute's conference (*Chemical Genomics/Chemogenomics*, Boston, MA, USA; 16 November 2001). The one-day conference presented an opportunity for academic and industry leaders to discuss their visions of chemical genomics, and comprised two sessions, *Functional Analysis* and *Lead Identification–Drug Development*, under the common banner of *High-Throughput Discovery of Disease Genes and Drugs*.

How are chemical genomics, chemogenomics and chemical genetics defined? Unfortunately, these terms are used interchangeably. In practice, most people define chemical genetics as 'the use of small synthetic molecules in a genetic approach to identify key genes involved in specific biological pathways'. Chemical genomics and chemogenomics refer to the use of small synthetic molecules that are highly specific for defined protein targets, for gene function analysis and to discover new drug leads. For gene function analysis, chemical genomics/chemogenomics can be applied to multiple members of a single gene family or to a disparate collection of disease-associated proteins (e.g. unrelated genes identified by genetics or transcription profiling).

Functional analysis

Brent R. Stockwell (Whitehead Institute, Cambridge, MA, USA) and Steven Zheng (Washington University School of Medicine, St Louis, MO, USA) described the use of exogenous target-specific ligands in a genetic approach to dissect biological processes; both used the rapamycin–TOR system (TOR being the target of rapamycin) as a model for this approach. Their studies were conducted in yeast, taking advantage of the rapid progress made in the yeast genome deletion project, which has >6000 gene knockouts. Sensitivity screens, consisting of monitoring yeast-cell growth in the presence of rapamycin, were conducted

 1359-6446/02/$ – see front matter PII: S1359-6446(02)02185-2

using various knockout strains; these revealed that up to 6% of yeast genes are associated with the TOR pathway. A second strategy discussed by Stockwell uses synthetic lethal screens, which aim to identify compounds that induce cell death in a cell-type specific manner. Using such a strategy, compounds were identified that selectively target tumor cells with minimal effects on primary fibroblasts.

Cancer and diabetes

Stuart L. Schreiber (Harvard University, Cambridge, MA, USA) spoke of his aims to integrate chemical diversity, molecular cell biology and genomics information to gain an insight into complicated diseases, such as cancer and type 2 diabetes. In his work, modifier screens – also referred to as phenotypic screens – are used, which evaluate the ability of a synthetic compound to enhance or modify a specific cellular phenotype. Uretupamine was discovered as an inhibitor of the transcription factor Ure2p, which is a crucial component of the yeast nitrogen discrimination pathway. Uretupamine was discovered by incubating Ure2p with high-density microarray slides printed with small molecules and looking for spots where the protein–ligand interaction occurred. This compound was later found to enhance the expression of an effector protein involved in signaling downstream from Ure 2P. In the area of type 2 diabetes, another compound was found to reverse the growth inhibitory effect of rapamycin in the presence of 10 mM glucose. In a recent communication, this group reported that the target for SMIR4 is ROT1p. These are examples where phenotypic screens were used to identify compounds, which in turn are used to unravel the biological processes responsible for this phenotype.

Mark Vacella (Cellular Genomics, Branford, CT, USA) introduced the company's integrated 'Gene to Screen to Lead' technology platform, emphasizing the P-target™ and the P-inhibitor™ technologies. Cellular Genomics' goal is to apply these tools to identify direct natural polypeptide substrates for orphan kinases (in the case of P-target™) and to validate novel kinases as drug discovery targets using synthetic modulators of the kinase activity. Both products are based on an interesting concept, the Analog Sensitive Kinase Alleles (ASKA), which involves genetic modification of the ATP-binding pocket of any kinase so that it can accept a selective ATP analogue. Cellular Genomics is also using P-inhibitor™ as an alternative to kinase-gene targeting in mice when this strategy results in an embryonic lethal phenotype. In an ASKA knockin mouse, the analysis of kinase function could be investigated using a specific small-molecule inhibitor targeting the kinase of interest.

Protein display

Two approaches to protein–DNA/RNA displays were discussed. Protein display enables rapid identification and cloning of effector proteins for any molecule of interest (e.g. small compound, DNA or polypeptide). The binding profile of an entire proteome against a lead compound could thus be assessed in a single experiment. Michael McPherson (Phylos, Lexington, MA, USA) described the company's proprietary mRNA display technology, PROfusion™, which involves the puromycin-mediated covalent linkage of polypeptides to the mRNA from which they are translated. These RNA–protein complexes are then exposed to the immobilized ligand. The bound polypeptides are enriched, upon several binding and reverse-transcription-PCR amplification cycles, culminating in the cloning and identification of the target gene.

Min Li (Johns Hopkins University School of Medicine, Baltimore, MD, USA) discussed the use of a proprietary DNA–protein display technology (ProCode™) to study chemical–protein interactions. In this case, newly synthesized proteins bind covalently to a DNA linkage site present on the recombinant ProCode™ vector from which they were translated. The proof-of-concept experiments for both PROfusion™ and ProCode™ were successfully conducted using FK506–FKBP12 as a model system. Both groups identified additional FK506 effector proteins, so it will be interesting to compare the identity of these proteins from each group.

This session was concluded with a presentation by Rachel Kindt (Exelixis, South San Francisco, CA, USA) who covered a lower throughput approach to small-molecule mediated target discovery. Exelixis' mechanism of action program was developed with the idea of using small model organisms as tools to identify protein targets and/or mechanism of action of approved compounds with established therapeutic activity but no known cognate target. The presentation focused mainly on studies performed using *Caenorhabditis elegans*. Exelixis has access to >18,000 knockout cell lines with transposon insertions in annotated positions within the genome, 1000 deletion mutants covering the entire genome and >6000 over-expresser cell lines, which are a valuable tool to identify genes implicated in the observed compound-induced phenotype.

Lead identification–drug development

The *Lead Identification–Drug Development* session kicked-off with an excellent overview, by Mark Murcko, of the chemogenomics program at Vertex (Cambridge, MA, USA). Murcko described the integrated drug design approach that the company is developing to build its chemogenomics platform, which aims at collating data obtained by computational and medicinal chemistry, structural biology and bioinformatics. Integration of this data enables the simultaneous design of drugs against multiple members of a protein family, a process referred to as 'scaffold morphing'.

Two protein families that are actively pursued at Vertex are the kinase and protease families with representative members being p38 mitogen-activated protein (MAP) kinase and interleukin I β-converting enzyme (ICE), respectively.

Scott Selleck (Arizona Cancer Center, Tucson, AZ, USA) related his institution's successful experience with Telik's Target Related Affinity Profiling or the TRAP™ system. Whereas most companies are expanding their HTS capabilities to hundreds of thousands of screening events per day, Telik's TRAP™ system offers the possibility of identifying low to submicromolar active molecules by screening as few as 200 compounds per target. The TRAP™ database comprises millions of data points that describe a library of small molecules; this information is used in conjunction with the classical physicochemical descriptors to identify a subset of compounds to be used in the first round of screening. Selleck concluded by describing some of the successes his institute has had with this screening approach, which include the discovery of ligands for kinases and integrins and even whole animal screens (*Drosophila*).

David O'Reilly introduced Inconix's (Mountain View, CA, USA) chemoinformatics tool, DrugMATRIX™, which can integrate a large body of information generated about known drugs to predict the efficacy and ADME properties of hit, lead and preclinical candidate molecules.

Target validation

Satish Jindal (NeoGenesis Pharmaceuticals, Cambridge, MA, USA) discussed the use of small-molecule ligands for target validation and initial lead discovery. NeoGenesis has built an ultra-high-throughput ligand identification system, ALIS™, which employs affinity selection using proteins in solution and highly complex and diverse compound libraries. The ALIS™ platform is a virtually hands-free, scalable and generic assayless technology, which is capable of querying any protein against 500,000 compounds per day with modest protein consumption and no functional information. Jindal reviewed several successful applications of this technology to discover initial drug leads against previously intractable protein targets. NeoGenesis is engaged in applying this technology to certain pharmaceutically and disease-relevant gene families.

Clearly the post-genomic world will require many novel approaches to advance our functional appreciation of newly discovered disease genes and proteins, and chemical genomics is an exciting tool for this task.